Engineering Applications of Artificial Intelligence 91 (2020) 103612

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Incorporating domain knowledge into reinforcement learning to expedite welding sequence optimization

Jesus Romero-Hdz [a,*], Baidya Nath Saha [b], Seiichiro Tstutsumi [a], Riccardo Fincato [a]

[a] Joining and Welding Research Institute, Osaka University, Japan
[b] Concordia University of Edmonton, Edmonton, Alberta, Canada

ARTICLE INFO

Keywords:
Welding sequence optimization
FEA based welding simulation
Reinforcement learning
Structural deformation
Residual stress
Artificial intelligence
Machine learning

ABSTRACT

Welding Sequence Optimization (WSO) is very effective to minimize the structural deformation, however selecting proper welding sequence leads to a combinatorial optimization problem. State-of-the-art algorithms could take more than one week to compute the best sequence for an assembly of eight weld beads which is unrealistic for the early stages of Product Delivery Process (PDP). In this article, we develop and implement a novel Reinforcement Q-learning algorithm for WSO where structural deformation is used to compute reward function. We utilize a thermo-mechanical Finite Element Analysis (FEA) to predict deformation. The exploration–exploitation dilemma has been tackled by domain knowledge driven ε-greedy algorithm into Q-RL which helps to expedite the WSO and we call this novel algorithm as DKQRL. We run welding simulation experiment using well-known Simufact® software on a typical widely used mounting bracket which contains eight welding beads. DKQRL allows the reduction of structural deformation up to ~71% and it substantially speeds up the computational time over Modified Lowest Cost Search (MLCS), Genetic Algorithm (GA), exhaustive search, and standard RL algorithm. Results of welding simulation demonstrate a reasonable agreement with real experiment in terms of structural deformation.

1. Introduction

Gas Metal Arc Welding (GMAW) is the most common technique for joining metal components and it has been preferred for its versatility, speed, and relative ease of robotic automation which is extensively used in automotive, shipbuilding, aerospace, construction, heavy and earth-moving equipment (Masubuchi, 1980; Islam et al., 2014). However, structural deformation due to welding is a serious concern for industry since it accrues various additional costs such as constraints in the design phase, extra operations, cost of quality, and overall capital expenditure (Goldak and Akhlaghi, 2005). WSO is highly cost effective which reduces welding structural deformation significantly. The ad hoc industry practice is to select the best sequence by experience and sometimes conduct a simplified design of experiments which typically leads to a sequence that generates considerably more structural deformation than the optimal one (Biswas et al., 2011). In order to get better welding sequence, it is needed to conduct innumerable real welding experiments which is very expensive and time consuming as well. To alleviate this problem, structural deformation yielded due to welding are predicted through a welding simulation software based on Finite Element Analysis (FEA) where thermo-mechanical models are commonly used and a reasonable solutions are achieved through FEA for numerous welding conditions and geometric configurations (Tikhomirov et al., 2005). There are three different FEA based models: (a) simplified: fast but less accurate, (b) thermo-mechanical: medium complexity but reasonable solutions, and (c) thermo-mechanical-metallurgical model: computationally very expensive and time consuming but highly accurate (Islam et al., 2014).

Selecting optimal welding sequence which ensues less deformation leads to a combinatorial optimization problem which is NP-hard by nature (Papadimitriou and Steiglitz, 1982). WSO can be mapped as a traveling salesman problem which is very popular in Operations Research (OR). Traveling Salesman (TS) problem can be described as given a list of cities and the distances between each pair of cities, discover the shortest possible route that visits each city exactly once and returns to the origin city. In the similar fashion, WSO can be described as given a list of welding seams to be placed along with all possible welding direction, find the best welding sequence which produces least structural deformation. The best welding sequence can be certainly found by executing the full factorial design of experiments. The total number of welding configurations for full factorial design can be counted by $N = n^r \times r!$, where n and r are the number of welding directions and beads (seams or segments) respectively. This number grows exponentially with the number of welding beads. For example, a complex weldment like an aero-engine assembly, it might

* Corresponding author.
E-mail address: jaromero@posgrado.cidesi.edu.mx (J. Romero-Hdz).

https://doi.org/10.1016/j.engappai.2020.103612
Received 19 July 2019; Received in revised form 14 January 2020; Accepted 16 March 2020
Available online 25 March 2020
0952-1976/© 2020 Elsevier Ltd. All rights reserved.

have 52–64 weld segments (Jackson and Darlington, 2011). Hence, the full factorial design is not feasible for industrial applications and is often practically in-feasible even using FEA at the early stages of Product Delivery Process (PDP) (Romero-Hdz et al., 2017). In order to succeed in the rapidly evolving global manufacturing landscape, there is a pressing need to increase the competitiveness in the welding industry. Moreover quality and efficiency are main drivers. So, mega-trends such as Internet-of-Things (IoT), Industry 4.0 as well as the development and usage of advanced materials will be critical to future competitiveness (Lindgren, 2007). Process simulation enables the implementation of Artificial Intelligence (AI) and Machine Learning (ML) techniques, because usually a great amount of process output data is required and "time to market" and "Do It Right The First Time" are pushing the industry to exploit virtual tools. Fig. 1 illustrates the deformation problem and the AI framework where coupled FEA-AI virtual tool controls the amount of deformation instead of conducting real experiments to keep the Geometric Dimensioning and Tolerancing (GD&T) features within tolerance and ensure the assemblability.

In this research, we present a novel and efficient Reinforcement Learning (RL) algorithm for Welding Sequence Optimization (WSO) to improve the weld quality where structural deformation is used to compute the reward function. We utilized a thermo-mechanical FEA modeling to predict welding deformation. RL, in the context of AI, is a type of dynamic programming where the agent over time makes decisions to maximize its reward and minimize its penalty. In the welding context, the agent will be rewarded if the sequence (action) taken minimizes the overall structure deformation. The advantage of this approach to AI is that it allows an AI program to learn without a programmer spelling out how an agent should perform the task. An agent is allowed to learn in an interactive environment by trial and error using feedback from its own actions and experiences (Sutton and Barto, 1998). Unlike supervised learning where feedback provided to the agent is correct set of actions for performing a task or explicitly mention how to perform a task, RL learns without human intervention by using rewards and punishment as signals for positive and negative behavior, i.e., the agent receives rewards by performing correctly and penalties for performing incorrectly. On the other hand, while the goal in unsupervised learning is to find similarities and differences between data points, in reinforcement learning the goal is to find a good behavior, a suitable action model or a label for each particular situation that would maximize the long-term benefits (cumulative reward) that the agent receives. RL algorithm has been extensively used in different fields such as gaming, neuroscience, psychology, economics, engineering communications, engineering power systems, and robotics (Sutton and Barto, 1998).

Here, we make the following technical contributions:

(A) Lessen computational complexity of a combinatorial optimization problem: We incorporate domain knowledge into Q-learning algorithm to expedite the convergence and we call it "DKQRL". Proposed DKQRL algorithm commendably curtails the computational complexity over exhaustive search. We conducted the experiment on a mounting bracket which includes eight weld seams that can be applied in two welding directions. In this scenario, the total number of welding configurations for exhaustive search is 10,321,920. However, in this experiment the DKQRL converges after 40 welding configurations. The average execution time for each welding configuration using FEA simulation software is 30 min. Thus we reduce considerable amount of computational time. **(B) Solve the Exploration–Exploitation Dilemma of RL through Domain Knowledge:** RL algorithm can be accelerated as well as converged through suitable determination of exploration and exploitation at each stage of RL algorithm. According to the domain expert of welding, it is advisable to weld the bead near the Center of Mass (CM) first to lessen the structural deformation due to welding. In the first step of DKQRL algorithm, if the weld seam near the CM causes minimum deformation we allow more exploration than exploitation throughout the process. In addition, when one bead of each part of the system is welded, it enhances the rigidity of the whole system that allows more exploration since high rigidity resists structural deformation (Park and An, 2016). Thus domain knowledge controls the ratio of exploration and exploitation throughout RL algorithm and hence expedite WSO. **(C) Traveling Salesman Problem and Welding Sequence Optimization:** We cast the problem of WSO with TS problem. TS problem consists of visiting each city only once with minimum cost. Similarly, WSO consists of welding each seam only once. As soon as one bead is welded, we remove the bead from the set of allowable states and the corresponding welding directions from the set of allowable actions. Mapping WSO with TS facilitates implementing RL in WSO and provides a realistic solution for the combinatorial optimization algorithm. **(D) State-of-the-art performance:** We conducted the simulation experiment of Gas Metal Arc Welding (GMAW) through the well-known welding simulation software Simufact®. The average execution time for each welding configuration took 30 min using a workstation with two Intel® Xeon® @2.40 GHz, 48G GB of RAM and 4 GB of dedicated video memory. The study case is defined as a typical mounting bracket which is widely used in telescopic jib (Derlukiewicz and Przybyłek, 2008) and automotive industries (Subbiah et al., 2011; Romeo et al., 2016). We validated the simulation results through real floor-shop welding experiment. Results demonstrated a high agreement between the result of simulation and real experiment in terms structural deformation. Experimental results demonstrated that best welding sequence can reduce significantly the amount of structural deformation (~71%) over worst sequence. DKQRL based approach substantially speeds up the computational time over standard RL, Genetic Algorithm (GA) and exhaustive search.

The organization of the paper is as follows. Section 2 presents literature review. Proposed domain knowledge driven reinforcement learning algorithm is presented in Section 3. Results are demonstrated in Section 4. Section 5 concludes this work. Relevant references are listed at the end of the paper.

2. Literature review

The literature review is organized into three parts. First, we summarize state-of-the-art optimization techniques implemented in fields related to welding that can be used for further research in WSO such as manufacturing process parameters optimization, mechanical and structural design optimization, Second, we describe Q-learning and RL approaches. Subsequently, we illustrate the domain knowledge for WSO.

2.1. State-of-the-art optimization techniques in related fields

Different optimization algorithms such as Genetic Algorithm (GA) (Islam et al., 2014; Romero-Hdz et al., 2016a, 2017a) , Particle Swarm Optimization (PSO) (Wang et al., 2015), Graph Search (Romero-Hdz et al., 2016b), Artificial Neural Networks (ANN) (Fukuda and Yoshikawa, 1990) have been proposed for WSO. Other popular methods for WSO include Joint Rigidity Method (Park and An, 2016), Surrogate Models (Voutchkov et al., 2005) and the use of generalized guidelines.

Yildiz et al. (2019a) present a comparative study on the application of ten recent meta-heuristic approaches for mechanical design optimization. In the study they solved six challenging mechanical problems of mixed variables type. The compared algorithms are: artificial bee colony (ABC), particle swarm optimization (PSO), moth-flame optimization (MFO), ant-lion optimizer (ALO), water cycle algorithm (WCA), evaporation rate WCA (ER-WCA), gray wolf optimizer (GWO), mine blast algorithm (MBA), whale optimization algorithm (WOA), and Salp swarm algorithm (SSA). Solved problems involve multiple objectives and mixed variables (continuous, integer and discrete), in addition to various nonlinear constraints on kinematic conditions, manufacturing requirement and performance operating. In terms of the

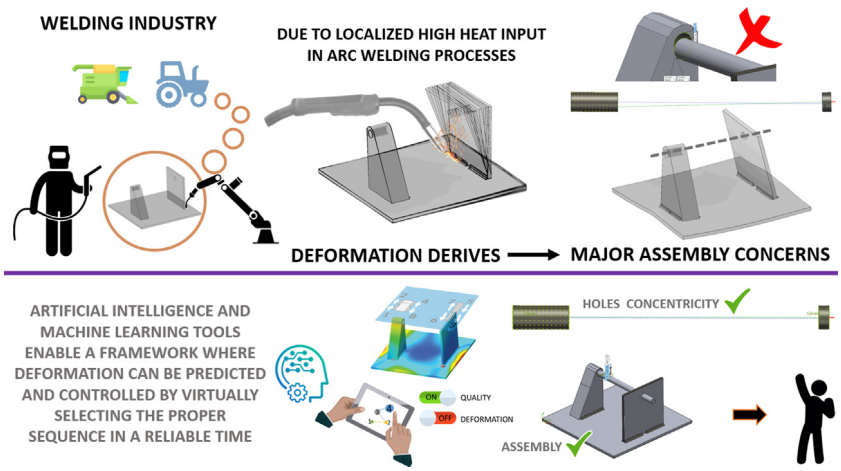


Fig. 1. Deformation problem, impact and WSO optimization framework as solution.

solution quality and the robustness, MFO was better than the other used algorithms.

Yildiz and Yildiz (2017) introduced a newly developed moth-flame optimization algorithm (MFO) for solving optimization problems in manufacturing industry. The optimization aim is to maximize the profit rate for multi-tool milling operations considering difficult constraints. The optimum manufacturing parameters is very important in order to develop the quality of the products and to enhance the total profit rate. According to Literature (Yildiz et al., 2019a; Yildiz and Yildiz, 2017) is more effective than many other population based-algorithms, in terms of welding, every single bead can be model as moth, so the population matrix can be a welding sequence to be conducted. For all the moths, it is assumed that there is a set of proper fitness values, in this case the deformation. In the MFO, both moths and flames are the solution. The moths are the real search agents who move around the search area, while the flames represent the best position of the moths obtained so far. In other words, flames can be seen as flags or pins thrown by moths when looking for a search space. For this reason, each moth searches around a flag (flame) and updates it to find a better solution.

Yildiz et al. (2019b) and Yıldız et al. (2019) have solved manufacturing problems by implementing different algorithms like the Harris hawks optimization algorithm (HHO), the grasshopper optimization algorithm (GOA), the multi-verse optimization algorithm (MVO), cuckoo search (Yildiz, 2013a), particle swarm optimization (Gökdağ and Yildiz, 2012), Hybrid Taguchi-Harmony Search (Yildiz and Ozturk, 2010) and artificial bee colony algorithm (Yildiz, 2013b). This work address a problem in grinding process to provide optimal surface quality by computing the optimal process parameters. The results reveal that the HHO, the MVO and the GOA outperforms the GA which have been one of the best in WSO.

Yildiz and Yildiz (2019) and Yildiz and Solanki (2012) have introduced several algorithms to solve shape optimization problems in the automotive industry. The algorithms are the Harris hawks optimization algorithm (the HHO), the Salp swarm algorithm (SSA), the grasshopper optimization algorithm(GOA), and the dragonfly algorithm (DA).

Pholdee et al. (2017) proposed the many-objective hybrid real-code population-based incremental learning and differential evolution algorithm (MnRPBILDE) based on the concept of objective function space reduction. The method is then implemented on real engineering design problems. The topology, shape and sizing design of a simplified automotive floor-frame structure are formulated and used as test problems.

2.2. Q-learning (reinforcement learning) for WSO

Okumoto and Ogawa (2007) developed a Q-learning (simplified RL algorithm) to optimize the welding route of an automatic machine. The machine consists of a simple truck system that can move only in one direction until a force is detected. A time-dependent fitness function is proposed in this work where this type of machines are moved from one joint to another manually by the welder. Since, the total number of possible execution combinations grow exponentially with the number of welding seams, a bad decision can increase the labor hours which have considerable effect on the cost and lead time. This method used ε-greedy based selection method which stochastically produces a reduced reward to avoid the local optima.

Romero-Hdz et al. (2018) implemented RL for WSO. This algorithm estimates the value function $Q(s,a)$ by repeating the action through trial and error in the environment. $Q(s,a)$ defines as the expectation of a gain when the agent takes the most suitable action in a state s. The most suitable action is expressed as the action a in state s, for which the value of $Q(s,a)$ becomes greatest among all the permissible actions in state s. The Q value is updated by the following Eq. (1).

$$Q(s,a) \leftarrow Q(s,a) + \alpha\{r(s,a) + \gamma \max_{a' \in A(s')} Q(s',a') - Q(s,a)\} \tag{1}$$

Where, $Q(s,a)$ is the value of action a in state s, $Q(s',a')$ is the value of action a' at state s' after transition, α is the learning rate ($0 < \alpha < 1$), and γ is the discount rate ($0 < \gamma < 1$). There are a number of selection methods for solving the exploration–exploitation dilemma available in the literature. Among them, we tailored very popular ε-greedy method for WSO in this study. This method stochastically accepts a reduced reward to circumvent the local minima. The ε-greedy chooses an action a in state s for which the value of $Q(s,a)$ is maximum at probability $(1-\varepsilon), 0 < \varepsilon < 1$, as illustrated in Eq. (2).

$$p(s,a) = \begin{cases} 1-\varepsilon & \text{when } Q(s,a) \text{ is maximum} \\ \varepsilon & \text{elsewhere} \end{cases} \tag{2}$$

They solved the exploration–exploitation dilemma by selecting the value of ε as 0.2 (generate a random number between 0 and 1 and if it is less than or equal to 0.2 , execute exploration, otherwise, perform exploitation). They selected the weld seam with a particular welding direction which gives the minimum of the maximum structural deformation for exploitation and on the other hand selected the second best weld seam for exploration. In this experiment, the robot or human is regarded as agent, the actions of the agent are the available weld seams that can be placed into the work piece along with the direction of the welding, the state is defined as the set of actions that had already executed. The reward is defined as the inverse of the maximum structural deformation. The most suitable action is defined as welding the weld seam along a particular direction which accrues minimum of the maximum structural deformation.

2.3. Domain knowledge for WSO

Fukuda and Yoshikawa (1990) discussed that welding sequence is determined based on the following heuristic knowledge to reduce the welding distortion, (a) weld from weld lines with greater restraint and shrinkage; (b) weld alternative weld lines in a member which are symmetrical to the neutral axis: weld on both sides alternating the sequence with the same proportion of weld when the joint is symmetric for example a V groove on both sides of the plate (Electric, 2018; Beardsley, 2018; Fukuda and Yoshikawa, 1990); (c) weld the closest weld lines first; (d) weld to avoid abrupt cooling at the ends of crossing weld lines; (e) weld symmetrically structural wise; (f) weld from the members nearest from the center of a structure: distortion is reduced by providing less leverage for the shrinkage forces (Okumoto and Ogawa, 2007); and (g) weld so as not to produce the unweldable parts after fabrication.

Literature (Electric, 2018; Beardsley, 2018) also reveals that the following guidelines could significantly improve the welding quality: (i) *Select a proper welding sequence:* the aim is to modify the thermal pattern in the weldment (Masubuchi, 1980); traditionally the sequence is selected by experience. Recently, different approaches using AI and ML have been proposed (Romero-Hdz et al., 2016a, 2018, 2017,a, 2016b); (ii) *Fewer weld passes:* in case of sizes over 10 mm in fillets or grooves a multi-pass welding is required; so, it is better few big passes than greater number of small passes with small electrodes; (iii) *Stress Relieving:* several conventional methods such as heat treatment, shot peening, modification of the structural configuration, usage of Low Temperature Transformation (LTT) wires (Romero-Hdz et al., 2017b) and thermal tensioning techniques are used; (iv) *Intermittent welding:* placing segments in a noncontinuous way across the joint; (v) *Clamping:* this is one of the most widely used methods, it consists of fixing the parts by a contact mechanism. New designs consider a water-cooled jig; (vi) *Use the back step welding technique:* it is a well known technique where basically the deposition is from right to left but each segment is deposited from right to left starting away from the heated zone; (vii) *Minimize welding time:* the energy per unit length or heat input can be decreased by setting up a different set of parameter, where the speed is an important factor. It is clear that the bead size can be slightly affected but finishing the welding cycle quickly has been proved to be effective; (viii) *Avoid over welding:* this item refers to control sizing and length of beads; (ix) *Presetting the parts:* this technique refers to tack the parts with a certain deviation on the opposite way where the shrinkage will pull the part out.

Kim et al. (2005) showed that there must be a certain period of time (delay) between welding operations associated with weld lines near each other to avoid heat-caused distortion.

3. Methodology

Here we present a novel RL algorithm where domain knowledge in the field of welding discussed in the previous section has been incorporated for accelerating WSO by solving the exploration–exploitation dilemma through adapting ε-greedy algorithm. In this section, we first outline the optimization framework, then we describe the implementation of the welding domain knowledge exploited in this study and lastly, the proposed DKQRL tailored for WSO is detailed.

3.1. DKQRL Framework

Fig. 2 illustrates the framework of DKQRL. Consider that an agent subsists in a world or environment, E, which can be defined by a set of states, S. First, the agent gets observations from the environment E, then the agent solves the exploration–exploitation dilemma for exploring new states and gets new information while maximizing its reward at the same time by taking action based on the information it receives and trusts it. Then the agent chooses an action a, consequently receives an immediate reward r and go to next state. In the similar fashion, the agent acts repeatedly, so that it receives maximum reward from an initial state to a goal state. RL enables to handle moderate dynamic change in the environment through learning by repeating and evaluating the action (Sutton and Barto, 1998).

3.2. Domain knowledge implementation

In this research, we have analyzed in detail the guidelines presented in the Section 2. As a result, we incorporate two general guidelines that can be used in any geometry without compromising the original design, quality of the bead completed or adding additional processes. First, we compute the Center of Mass (CM) of the whole system and then calculate the euclidean distance of the nearest and farthest point of each of the welding bead from CM. We resolve the issue of exploration–exploitation dilemma of RL by placing the weld bead near and around the CM first to minimize the welding deformation. This guideline takes into account the recommendations of Literature (Fukuda and Yoshikawa, 1990) described in Section 2 as (a), (c) and (f).

Second, we draw a thermal bounding box around the bead based on the HAZ (Heat Affected Zone) distance. As soon as the first bead is welded, we avoid to select the weld bead located within the thermal box region or HAZ distance of previous welded bead. During heating and cooling cycle of welding operation, many factors affect shrinkage of the metal and lead to distortion, such as physical and mechanical properties of the metal which change as heat is applied. For an example, as the temperature of the weld area increases, yield strength, elasticity, and thermal conductivity of the steel plate decrease, while thermal expansion and specific heat increase. These changes, in turn, affect heat flow and uniformity of heat distribution which affect structural deformation. Rocha et al. (2018) measure the temperature profiles at different distances from the weld line using infrared thermography. Temperature at 10 mm of distance from the weld line achieves around 425 °C and the phase transition does not occur. We draw a thermal box of 30 mm region across all the previously welded bead along all directions which is consistent with Kim et al. (2002) and suggest not to weld in this region as illustrated in Fig. 3.

This guideline takes into account the recommendations of Literature (Fukuda and Yoshikawa, 1990) described in Section 2 as (d), recommendations of Literature (Electric, 2018) as (vi), and the one by Literature (Kim et al., 2005). In Literature (Fukuda and Yoshikawa, 1990), guidelines (b) and (e) refer to symmetric beads location, something that is not generalized for all geometries. Similarly, in Literature (Electric, 2018), (ii) was not implemented because this study case is not multi-pass welding, (iii) are additional processes prior, during or post welding, (iv) is related to design change and structural analysis is required to validate the impact on the performance of the part,(v) fixtures usually increases the cost of the project implementation and prototypes are often build without tooling, (vii) changes on the parameters might affect the bead finish so the optimization of them should run separately, (viii) is also related to parameters and for this study case we use a robot, so bead size is controlled.

Finally suggestions (i) in Literature (Electric, 2018) and (g) in Literature (Fukuda and Yoshikawa, 1990) are the main goal of WSO which has been adopted in the approach proposed in this study. All other guidelines are not compatible with the experimental setup of this proposed study and hence have not been included in this study.

3.3. Domain knowledge based q-learning (DKQRL)

We adjust the value of ε of ε-greedy algorithm for WSO by incorporating domain knowledge which expedites Q-learning. Algorithm 2 demonstrates the domain knowledge based reinforcement Q-Learning (DKQRL) for WSO. Unlike the conventional ε-greedy algorithm where the value of ε is fixed, in this modified ε-greedy algorithm, we increase

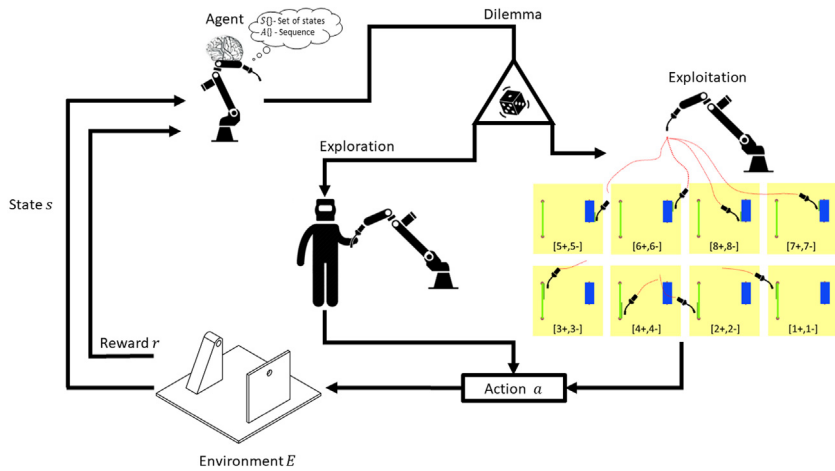
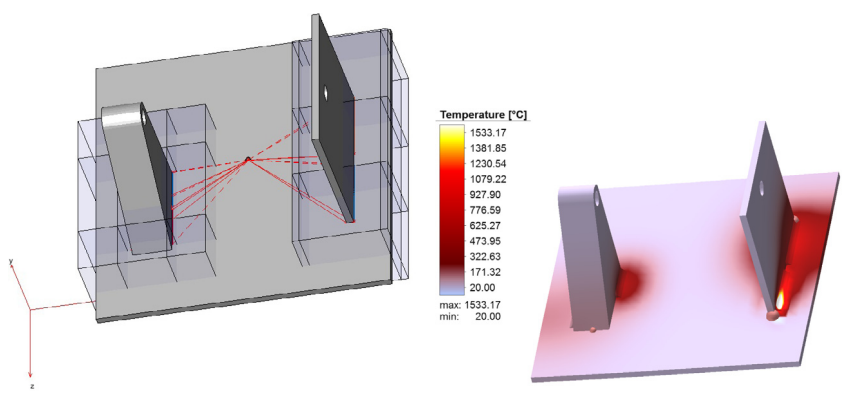


Fig. 2. RL with domain knowledge (DKQRL) framework for WSO.



(a) Representation of the thermal bounding boxes and distances from center of the mass to each bead.

(b) Thermal field of welding simulation.

Fig. 3. Representation of the implemented domain knowledge guidelines.

Fig. 4. Different mounting brackets available in the market similar to our study case.

the value of ε at each stage of RL algorithm, if the suitable action chosen by exploration agrees with the action chosen by exploitation which increases the probability of choosing the exploration over exploitation as demonstrated in Algorithm 1. Since FEA based welding simulation is computationally expensive and WSO is a combinatorial optimization problem, hence exploration reduces the computational complexity significantly over exploitation.

The advantages of exploiting domain knowledge are two-fold: (i) domain knowledge facilitates to select those actions which reduces the welding structural deformation through exploration; (ii) domain knowledge resolves exploration–exploitation dilemma by increasing the probability of ε, which allows more exploration than exploitation, and thus accelerates WSO by selecting those actions responsible for minimizing structural deformation without conducting all permissible actions through exploitation.

4. Experimental results and discussions

In this section, first we illustrate the study case. Then, we discuss FEA based simulation experiment conducted for welding deformation prediction. Subsequently, we illustrate the results of the FEA for the best and worst sequence found by the proposed DKQRL algorithm. After that, we demonstrate the effects of welding sequence on WSO. Next, we demonstrate a comparative study among Modified

Algorithm 1: Pseudocode of Exploration–Exploitation Dilemma tackled by DKQRL for WSO

Input: the value of ε;
Observe the current state S;
Generate a random number, say r;
if $r \leq \varepsilon$ **then**
  Exploration: Choose an action a (select the weld bead and the weld direction which has the minimum distance from the CM and does not overlap with any of the thermal box region of the seams that are already welded); $a = \text{argmin}_{a' \in A'}[Q(s', a')]$;
  /* $Q(s', a') = \text{argmax}_{a' \in A'} \frac{1}{d(a', C)}$
  such that $A(a') \cap A(T) = \varnothing$ (Null set),
  Where $d(a', C)$ represents the distance between the weld bead associated with a' and CM (Center of mass) of the system, $A(x)$ represents the area of the parameter x and T represents the thermal box regions; */
  Exploitation: Choose an action a (select the weld bead and the weld direction) from the set of allowable actions (available welding seams that are not welded yet and with all possible welding directions, A') which provides minimum of the maximum structural deformation; $a = \text{argmin}_{a' \in A'}[Q(s', a')]$;
  /* $Q(s', a') = \text{argmax}_{a' \in A'} \frac{1}{deformation(s')}$, minimum of the maximum structural deformation; */
else
end
Execute action a;
Update Q-value with $Q(s, a) \leftarrow min_{a'}[Q(s', a')]$ and remove a and $-a$ from A';
Update state-transition model $P(s'|s, a)$;
Store $S = \{s\}$ and $A = \{A, a\}$;
Return S and A;

Algorithm 2: Pseudocode of Domain Knowledge based Reinforcement Q -Learning (DKQRL) algorithm for WSO

Initialize $Q(s, a) = 0,\ S = \{\},\ A = \{\},\ s = \{\},\ A' = \{+1, -1, +2, -2, \ldots, +m, -m\}$; where m is the number of weld beads, '+' and '−' represents two different weld directions.
Initialize a state-transition model $P(s'|s, a)$
$P(s'|s, a) = 1/deformation(s')$;
$\varepsilon = \varepsilon_0$;
Result: Weld sequences having minimum of the maximum welding deformation
while *learning do not end i. e., do not finish all the cycles or iterations* **do**
  a) Compute the Center of Mass (CM) of the whole system, and then compute the linear distance between the CM and start and end of each of the weld bead;
  b) **Exploration: Choose** an action a' (Select the weld bead and the weld direction which has the minimum distance from the CM);
  c) Compute the thermal box of each of the weld bead (30 mm in all the direction across a weld bead);
  d) **Exploitation: Choose** an action a'' (select a bead and the weld direction) from the set of allowable actions (available welding seams with all possible welding directions, A') which provides minimum of the maximum structural deformation;
  $a = \text{argmin}_{a' \in A'}[Q(s', a')]$;
  /* $Q(s', a') = \text{argmax}_{a' \in A'} \frac{1}{deformation(a')}$, minimum of the maximum structural deformation; */
  if $a' == a''$ **then**
    $\varepsilon = min(\varepsilon_0 + \lambda, \xi)$; $0 < \varepsilon,\ \lambda,\ \xi < 1$;
  else
    $\varepsilon = \varepsilon_0$; $0 < \varepsilon_0 < 1$;
  while *we do not reach a goal i. e., all the seams are not welded* **do**
    if *at least one seam of each part is welded* **then**
      Solve the exploration–exploitation dilemma by executing Algorithm 1 with $\varepsilon = min(\varepsilon + \lambda, \xi)$;
    else
      Solve the Exploration–Exploitation Dilemma by executing Algorithm 1 with $\varepsilon = \varepsilon$;
Return A;

Lowest Cost Search (MLCS) (Romero-Hdz et al., 2016a,b), single objective Genetic Algorithm (SOF-GA) (Romero-Hdz et al., 2016a), multi-objective Genetic Algorithm(MOF-GA) (Romero-Hdz et al., 2017a), Q-learning based RL method (Romero-Hdz et al., 2018), and proposed Domain Knowledge based Q-Reinforcement Learning (DKQRL) algorithm. Finally we validate the results of simulation with real experiment.

4.1. Mounting bracket used in this study

Mounting bracket as shown in Fig. 4 is widely used in telescopic jib (Derlukiewicz and Przybyłek, 2008), automotive industries (Subbiah et al., 2011), and cars (Romeo et al., 2016). Fig. 5 illustrates the engineering drawing including dimensional and welding specifications for this study. In the design there are two parts to be welded, from the common knowledge it is clear that the thinner part will concentrate more deformation that the one in triangle form. Since the tolerance of the true position between holes on both plates to be welded is 0.5 mm, the concentricity can be achieved by drilling the parts prior welding, which is a quite common strategy to reduce the cost in metal-fab operations (Rocha et al., 1995).

4.2. Deformation prediction by FEA

Welding deformation can be accurately predicted using FEA. Furthermore, the research on welding simulation has reached a high level of maturity and recently several types of software with a great Graphical User Interface (GUI) and features emerge to satisfy the industry needs. However, the industry stills on the way to fully implement welding simulation as part of their product delivery process(PDP) (Goldak, 2013). We conducted a simulation experiment of GMAW process using popular FEA welding software (Islam et al., 2014). We have incorporated domain knowledge into reinforcement Q-Learning for choosing

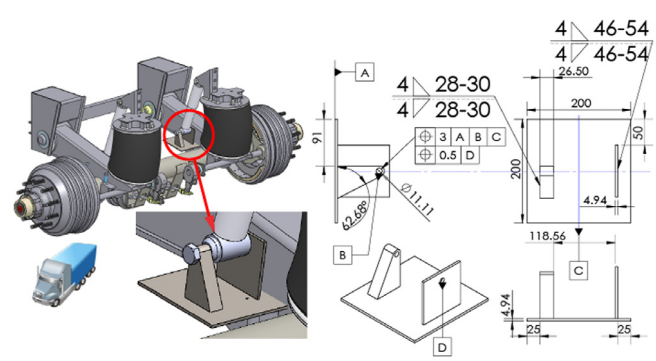


Fig. 5. Engineering drawing of the study case.

the best welding sequence having minimum structural deformation and we demonstrated the effects of welding sequence on the weld quality by analyzing the structural deformation caused by welding by choosing different welding sequences. We used the values of the parameters same ($\varepsilon_0 = 0.1,\ \lambda = 0.2$) throughout the experiment. We divided the welding bead into eight segments as shown in Fig. 5. We compared the results of proposed DKQRL with MLCS (Romero-Hdz et al., 2016a,b), single

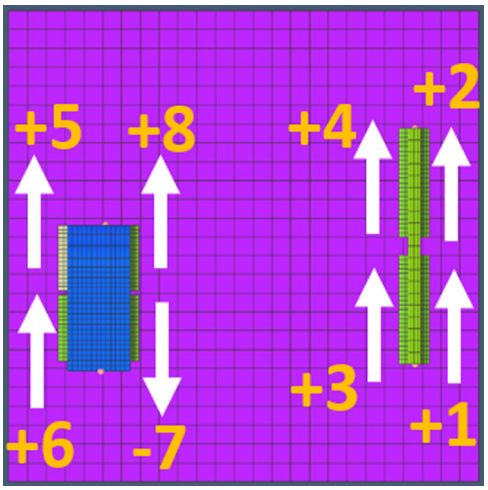


Fig. 6. Initial bead ID in the FEA modeling to track the sequence changes.

Table 1
Comparative analysis among different AI techniques in the case study.

AI Techniques	FEA runs	Computation time (h)	Deformation (mm)
MLCS	72	30	0.93
GA SOF	115	57.5	0.55
GA MOF	81	40.5	0.66
RL	30	18	1.41
DKQRL	40	7	0.61

objective GA (Romero-Hdz et al., 2016a), multi-objective GA (Romero-Hdz et al., 2017a), and reinforcement learning (Romero-Hdz et al., 2018) algorithm.

4.3. Discussions about the finite element analysis (FEA) results

Fig. 6 illustrates bead ID numbers, position and direction in the initial FEA modeling to track the sequence changes. The integer number represents the name of the bead, the position in the vector is the sequence, means order of placement starting from the left to right, and the sign is the direction of welding, for this case, looking at the part from the top view, positive means up, and negative down. The best and worst sequences found were $[+5, +8, -1, -7, +2, -4, -3, -6]$ and $[+3, -4, +7, -6, -5, +1, -8, +2]$ with maximum structural deformation of 0.61 mm and 2.76 mm respectively as shown in Fig. 7. We resolve the issue of exploration–exploitation dilemma by knowledge based ε-greedy algorithm, where we choose the value of ε as 0.1 at iteration 0. Then if the bead nearest to the CM generates minimum structural deformation, we increase the value of ε as 0.4. This indicates that the RL allows exploration 40% time and exploitation 60% time of the actions chosen. If the value of ε increases, it facilitates early convergence of RL algorithm by avoiding onerous FEA based computations of the structural deformation for all permissible actions and choosing only the best action among the set of all permissible actions by applying domain knowledge. When welding operation is performed over beads sequentially, the structural deformation at any stage cannot be determined as a function of previous stages and the reward function is not cumulative in nature, hence the value of the learning rate and discount factor are inapplicable in our application.

4.4. Effects of welding sequence on welding sequence optimization

Box plot of the structural deformation and effective stress values on the mounting bracket used in this study are shown in Figs. 8 and 9 respectively. If we consider the deformation value of the worst

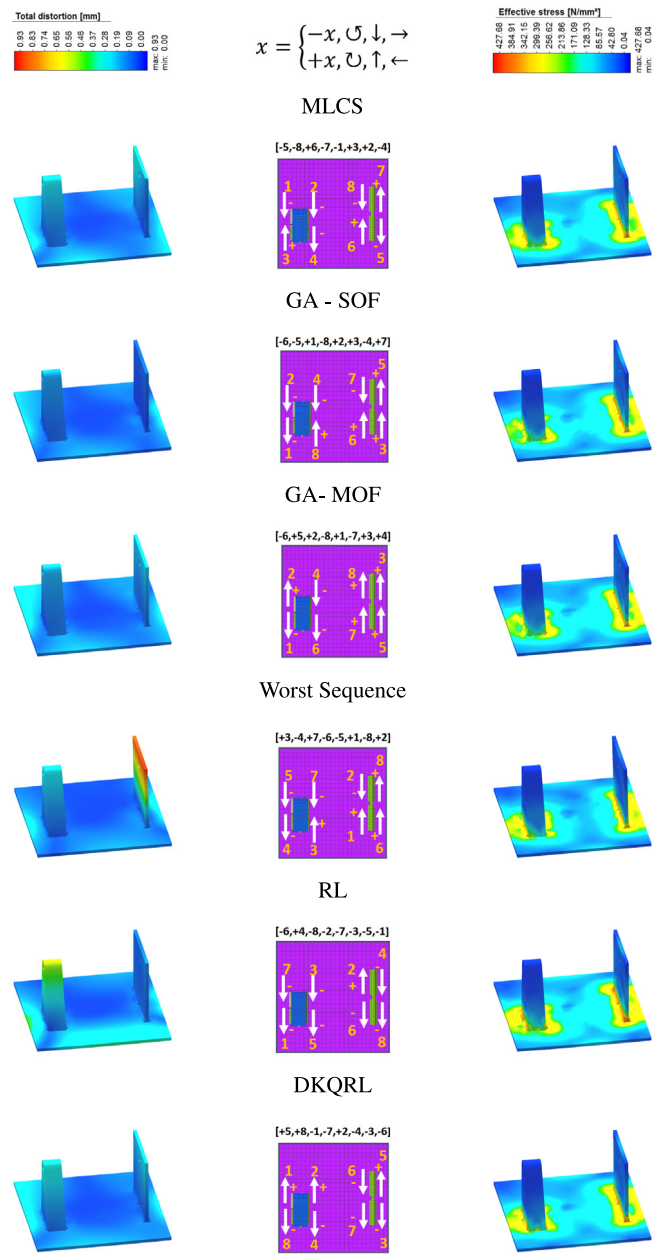


Fig. 7. Results of different AI techniques and worst sequence found in the simulation experiment.

sequence as 100%, RL algorithm reduces ∼71% maximum structural deformation over worst sequence (maximum structural deformation of best and worst sequence are 0.61 mm, and 2.76 mm respectively). It is evident that welding sequence has considerable effect on welding deformation, though, it has very little effect on effective stress which are consistent with Romero-Hdz et al. (2016a) and Romero-Hdz et al. (2017a).

4.5. Comparative analysis: Reinforcement learning vs. other AI techniques

The performance of the proposed DKQRL algorithm was compared with other available AI techniques used for WSO in the literature. For GA, we have used the same parameters reported in Romero-Hdz et al. (2016a) and Romero-Hdz et al. (2017a). Fig. 7 demonstrates the maximum structural deformation of the best sequence found by

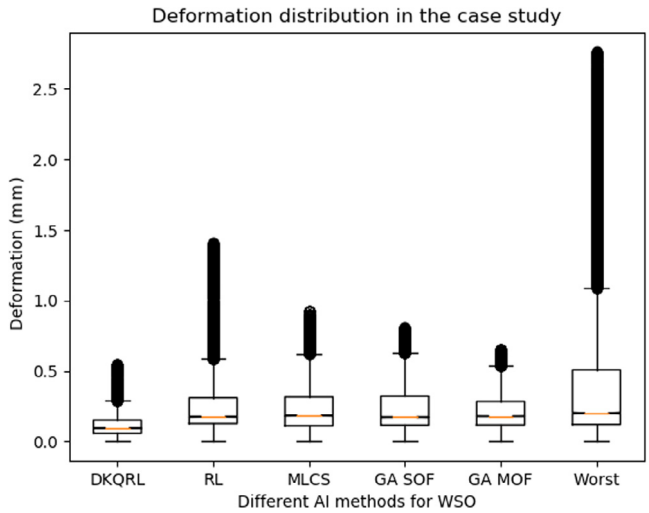


Fig. 8. Boxplot illustrating deformation of different AI techniques for WSO.

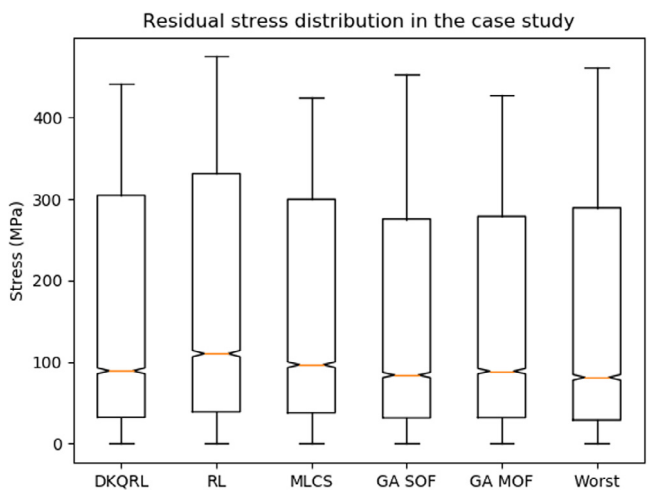


Fig. 9. Boxplot demonstrating residual stress of different AI techniques for WSO.

MLCS (Romero-Hdz et al., 2016a,b), the single objective GA (Romero-Hdz et al., 2016a), multi-objective GA (Romero-Hdz et al., 2017a) and RL (Romero-Hdz et al., 2018) methods. Table 1 demonstrates the comparative analysis among MLCS, single objective GA, multi-objective GA and RL method. Table 1 shows that RL produces a bit more structural deformation, however, RL converges much faster than GA. DKQRL converges after 40 FEA simulations whereas MLCS, single objective GA, multi-objective GA, and RL require 72, 115, 81, and 30 simulations respectively. Since it takes approximately 30 min for each welding simulation, DKQRL method takes 7 h whereas MLCS, single objective GA, multi-objective GA, and RL methods take 30, 57.5, 40.5, and 18 h respectively. Fig. 10 illustrates the search space (tree) for the best sequence excavated by the DKQRL algorithm.

4.6. Experimental validation of welding simulation

Fig. 11 demonstrates three dimensional (3D) view of the structural deformation side by side of both real and simulation experiment for the best sequence $[+5, +8, -1, -7, +2, -4, -3, -6]$. We illustrate 3D visualization of real experiment using Geomagic Control® software with a Creaform® Metrascan 210 C-Track 780 3D optical scanner. We used simufact® FEA software (Islam et al., 2014) for simulation experiment. Fig. 11 shows a reasonable agreement between real and simulation experiment in terms of structural deformation.

Results demonstrate that welding simulation predicts a bit higher deformation values than the real experiment. The difference between the real and simulation experiment occurs as a result of the cumulative error in the measurement process. Typically 3D scanner generates 0.085 mm of error and welding positioner accrues 0.1 mm of error. Deformation difference in mm measured by scanner and simulation are illustrated through boxplot as shown in Fig. 12. Fig. 12 demonstrates that the distribution of plate and base are same, however the distribution of triangle are different from others. It shows that the mean absolute difference of the deformation between the scanner and simulation at plate and base regions are 0.112 ± 0.084 mm where the difference is 0.1629 ± 0.1191 mm at triangle region. This demonstrates that results of simulation agree with the real experiment at plate and base regions more than the triangular region. The average deformation at plate, base and triangular regions found in real experiment are 0.1377 ± 0.1173 mm, 0.1430 ± 0.1103 mm, and 0.2763 ± 0.1676 mm respectively. It stipulates that the disagreement between the real and simulation experiment is higher at high deformation region. It points out that the simulation model needs suitable calibration at the high deformation region.

5. Conclusions and future work

Welding sequence optimization has considerable effect in structural deformation. In this study, the maximum structural deformation is exploited as the Q-function of the RL algorithm for WSO. RL significantly

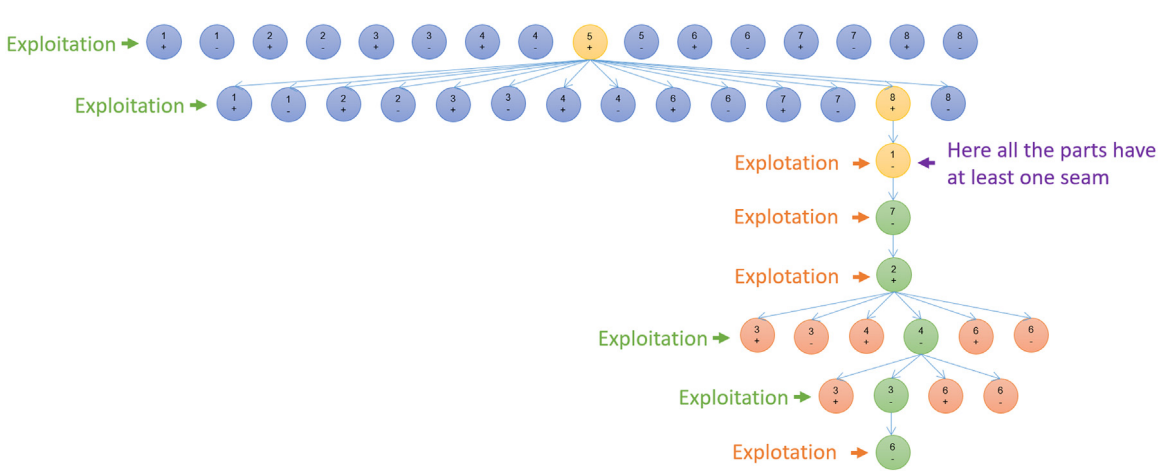


Fig. 10. Search tree explored by DKQRL algorithm.

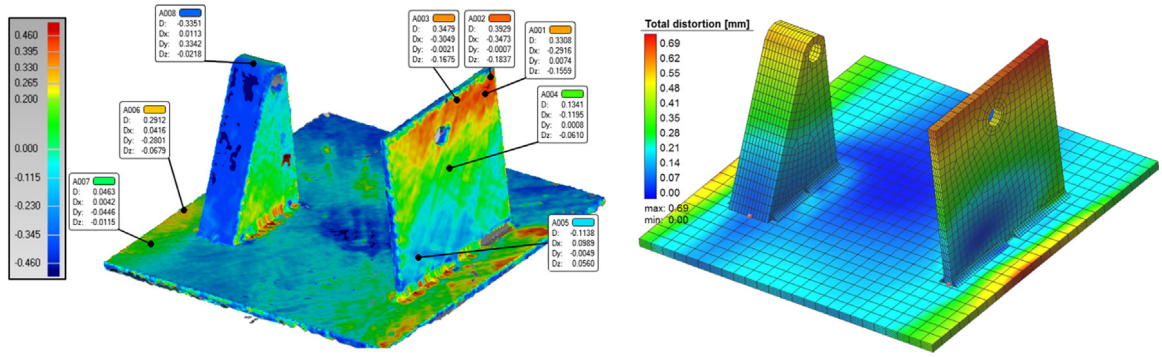


Fig. 11. Structural deformation comparison between real and simulated experiments.

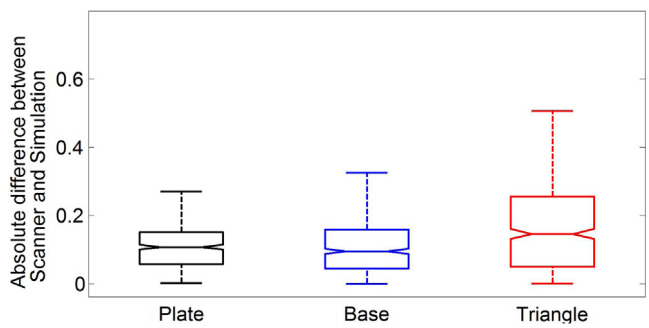


Fig. 12. Deformation difference (in mm) between real experiment and simulation.

reduces the search space over exhaustive search. We incorporated the domain knowledge and expedite the RL algorithm for WSO by resolving the exploration–exploitation dilemma. Welding simulation software was used to compute the structural deformation using FEA. Proposed DKQRL algorithm for WSO was compared with MLCS, single objective GA, multi-objective GA and RL algorithms. Proposed DKQRL based WSO method reduces the search space considerably over other AI techniques and thus RL finds the pseudo optimal welding sequence much faster than GA. Experiment of welding simulation was validated with real experiment.

This research unwraps different avenues for WSO research. In near future, we would like to develop a more robust multivariate reward function including structural deformation, residual stress, temperature and robot path time for WSO. Information of the deformation, residual stress and temperature at each stage in WSO require to be investigated for significant improvement of welding quality.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Baidya Nath Saha: Formal Analysis. **Seiichiro Tstutsumi:** Project administration. **Riccardo Fincato:** Writing - review & editing.

Acknowledgments

The authors gratefully acknowledge the support provided by Osaka University through the JWRI in Japan and The National Council of Science and Technology of Mexico through CIDESI and CIMAT.

References

Beardsley, K., 2018. Welding how-to's - problems - distortion. [cited 27.03.2018]. URL http://www.fabricatingandmetalworking.com/2013/02/how-to-control-the-warping-of-parts-in-thin-sheet-metal/.

Biswas, P., Kumar, D.A., Mandal, N.R., Mahapatra, M.M., 2011. A study on the effect of welding sequence in fabrication of large stiffened plate panels. J. Marine Sci. Appl. 10 (4), 429–436.

Derlukiewicz, D., Przybyłek, G., 2008. Chosen aspects of FEM strength analysis of telescopic jib mounted on mobile platform. In: 8th International Conference on Computer-Aided Engineering. Autom. Constr. 17 (3), 278–283.

Electric, L., 2018. Welding how-to's - problems - distortion. [cited 27.03.2018]. URL http://www.lincolnelectric.com/en-us/support/welding-how-to/Pages/weld-distortion-detail.aspx.

Fukuda, S., Yoshikawa, K., 1990. Determination of welding sequence: a neural net approach. Eng. Anal. Bound. Elem. 7 (2), 78–82.

Gökdağ, H., Yildiz, A., 2012. Structural damage detection using modal parameters and particle swarm optimization. Mater. Test. 54, 416–420. http://dx.doi.org/10.3139/120.110346.

Goldak, J., 2013. Web based simulation of welding and welded structures. In: CWA Conference 2013, p. 24.

Goldak, J., Akhlaghi, M., 2005. Computational Welding Mechanics. Springer US.

Islam, M., Buijk, A., Rais-Rohani, M., Motoyama, K., 2014. Simulation-based numerical optimization of arc welding process for reduced distortion in welded structures. Finite Elem. Anal. Des. 84 (Complete), 54–64.

Jackson, K., Darlington, R., 2011. Advanced engineering methods for assessing welding distortion in aero-engine assemblies. IOP Conf. Ser.: Mater. Sci. Eng. 26, 012018.

Kim, H.-J., Kim, Y.-D., Lee, D.-H., 2005. Scheduling for an arc-welding robot considering heat-caused distortion. J. Oper. Res. Soc. 56 (1), 39–50.

Kim, K.-Y., Kim, D.-W., Nnaji, B.O., 2002. Robot arc welding task sequencing using genetic algorithms. IIE Trans. 34 (10), 865–880.

Lindgren, L., 2007. Computational Welding Mechanics: Thermomechanical and Microstructural Simulations, first ed. CRC Press.

Masubuchi, K., 1980. Analysis of Welded Structures: Residual Stresses, Distortion, and Their Consequences. International Series on Materials Science and Technology, Pergamon Press Ltd..

Okumoto, Y., Ogawa, T., 2007. Optimization of welding route by automatic machine using reinforcement learning method. J. Japan Soc. Naval Archit. Ocean Eng. 5, 85–89.

Papadimitriou, C.H., Steiglitz, K., 1982. Combinatorial Optimization: Algorithms and Complexity. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Park, J.-U., An, G.B., 2016. Effect of welding sequence to minimize fillet welding distortion in a ship's small component fabrication using joint rigidity method. Proc. Inst. Mech. Eng. B 230 (4), 643–653.

Pholdee, N., Bureerat, S., Yildiz, A., 2017. Hybrid real-code population-based incremental learning and differential evolution for many-objective optimisation of an automotive floor-frame. Int. J. Veh. Des. 73, 20–53. http://dx.doi.org/10.1504/IJVD.2017.082578.

Rocha, E.J.F., Antonino, T.d.S., Guimaraes, et al., 2018. Modeling of the temperature field generated by the deposition of weld bead on a steel butt joint by FEM techniques and thermographic images. Mater. Res. 21.

Rocha, J., Vale, Z., Ramos, C., 1995. Sequencing operations for process planning. In: Practical Applications of Prolog. pp. 1–10.

Romeo, Ş.-I., Gheorghe, P., Vieru, I., Nicolae, V., Sebastian, P., 2016. Study of the dynamic behavior of a Car body for mounting the rear axle. In: Andreescu, C., Clenci, A. (Eds.), Proceedings of the European Automotive Congress EAEC-ESFA 2015. Springer International Publishing, Cham, pp. 349–356.

Romero-Hdz, J., Aranda, S., Toledo, G., Segura, J., Saha, B., 2016a. An elitism based genetic algorithm for welding sequence optimization to reduce deformation. J. Res. Comput. Sci. 121, 17–36.

Romero-Hdz, J., Saha, B.N., Toledo, G., 2016b. Welding sequence optimization through a modified lowest cost search algorithm. Comput. Sci. Eng. 6, 25–32.

Romero-Hdz, J., Saha, B., Toledo-Ramirez, G., Beltran-Bqz, D., 2017. Welding sequence optimization using artificial intelligence techniques, an overview. SSRG Int. J. Comput. Sci. Eng. 3, 90–95.

Romero-Hdz, J., Saha, B., Toledo-Ramirez, G., Lopez-Juarez, I., 2018. A reinforcement learning based approach for welding sequence optimization. In: Chen, S., Zhang, Y., Feng, Z. (Eds.), Transactions on Intelligent Welding Manufacturing. Springer Singapore, Singapore, pp. 33–45.

Romero-Hdz, J., Toledo, G., Saha, B., 2017a. Deformation and residual stress based multiobjective genetic algorithm for welding sequence optimization. J. Res. Comput. Sci. 132, 155–179.

Romero-Hdz, J., Tsutsumi, S., Fincato, R., Saha, B., 2017b. Influence of welding sequence on residual stress and deformation pattern using conventional and LTT wires. In: Welding Structure Symposium, Osaka, Japan, pp. 435–443.

Subbiah, S., Singh, O., Mohan, S.K., Jeyaraj, A.P., 2011. Effect of muffler mounting bracket designs on durability. Eng. Fail. Anal. 18 (3), 1094–1107.

Sutton, R.S., Barto, A.G., 1998. Introduction to Reinforcement Learning, first ed. MIT Press, Cambridge, MA, USA.

Tikhomirov, D., Rietman, B., Kose, K., Makkink, M., 2005. Computing welding distortion: Comparison of different industrially applicable methods. In: Sheet Metal 2005. In: Advanced Materials Research, vol. 6, Trans Tech Publications Ltd, pp. 195–202. http://dx.doi.org/10.4028/www.scientific.net/AMR.6-8.195.

Voutchkov, I., Keane, A., Bhaskar, A., Olsen, T.M., 2005. Weld sequence optimization: The use of surrogate models for solving sequential combinatorial problems. Comput. Methods Appl. Mech. Engrg. 194 (30), 3535–3551.

Wang, X.-W., Shi, Y.-P., Gu, X.-S., Ding, D.-Y., 2015. Partition mutation PSO for welding robot path optimization. In: Tarn, T.-J., Chen, S.-B., Chen, X.-Q. (Eds.), Robotic Welding, Intelligence and Automation. Springer International Publishing, Cham, pp. 77–86.

Yildiz, A.R., 2013a. Cuckoo search algorithm for the selection of optimal machining parameters in milling operations. Int. J. Adv. Manuf. Technol. 64, 55–61. http://dx.doi.org/10.1007/s00170-012-4013-7.

Yildiz, A.R., 2013b. A new hybrid artificial bee colony algorithm for robust optimal design and manufacturing. Appl. Soft Comput. 13 (5).

Yildiz, A.R., Abderazek, H., Mirjalili, S., 2019a. A comparative study of recent non-traditional methods for mechanical design optimization. Arch. Comput. Methods Eng..

Yildiz, A., Mirjalili, S., Sait, S., Li, X., 2019b. The harris hawks, grasshopper and multi-verse optimization algorithms for the selection of optimal machining parameters in manufacturing operations. Materialprufung 8, 1–15. http://dx.doi.org/10.3139/120.111377.

Yildiz, A.R., Ozturk, F., 2010. Hybrid taguchi-harmony search approach for shape optimization. In: Geem, Z.W. (Ed.), Recent Advances in Harmony Search Algorithm. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 89–98. http://dx.doi.org/10.1007/978-3-642-04317-8_8.

Yildiz, A.R., Solanki, K., 2012. Multi-objective optimization of vehicle crashworthiness using a new particle swarm based approach. Int. J. Adv. Manuf. Technol. 59 (1–4), 367–376. http://dx.doi.org/10.1007/s00170-011-3496-y.

Yildiz, B., Yildiz, A., 2017. Moth-flame optimization algorithm to determine optimal machining parameters in manufacturing processes. Mater. Test. 59, 425–429. http://dx.doi.org/10.3139/120.111024.

Yildiz, A., Yildiz, B., 2019. The harris hawks optimization algorithm, salp swarm algorithm, grasshopper optimization algorithm and dragonfly algorithm for structural design optimization of vehicle components. pp. 60–70.

Yıldız, A.R., Yıldız, B.S., Sait, S.M., Bureerat, S., Pholdee, N., 2019. A new hybrid Harris hawks-Nelder-mead optimization algorithm for solving design and manufacturing problems. Mater. Test. 61 (8), 735–743.